

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 542 863 B1**

(12)

## EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention  
of the grant of the patent:  
26.11.1997 Bulletin 1997/48

(51) Int Cl.<sup>6</sup>: **C12N 15/62**, C12N 15/63,  
C07K 1/00

(21) Application number: 91915080.5

(86) International application number:  
PCT/US91/05617

(22) Date of filing: 07.08.1991

(87) International publication number:  
WO 92/02550 (20.02.1992 Gazette 1992/05)

### (54) EXPRESSION OF RECOMBINANT POLYPEPTIDES WITH IMPROVED PURIFICATION

EXPRESSION REKOMBINANTER POLYPEPTIDE MIT VERBESSERTER MÖGLICHKEIT ZUR  
REINIGUNG

EXPRESSION DE POLYPEPTIDES RECOMBINES AVEC PURIFICATION AMELIOREE

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

- HILLIKER, Sandra  
Riverside, CA 92503 (US)
- WILLETT, W., Scott  
San Francisco, CA 94117 (US)

(30) Priority: 07.08.1990 US 564259

(43) Date of publication of application:  
26.05.1993 Bulletin 1993/21

(74) Representative: Goldin, Douglas Michael et al  
J.A. KEMP & CO.  
14 South Square  
Gray's Inn  
London WC1R 5LX (GB)

(73) Proprietor: SCIOS INC.  
Mountain View, CA 94043 (US)

(72) Inventors:  
• TARNOWSKI, S., Joseph  
Sunnyvale, CA 94086 (US)

(56) References cited:  
EP-A- 0 089 626                      US-A- 4 431 739  
US-A- 4 532 207                      US-A- 4 743 679  
US-A- 4 987 070

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 542 863 B1

**Description**Technical Field

5 This invention relates to the molecular biology of hormonal peptide actors and recombinant DNA technology. More specifically, this invention relates to atrial natriuretic peptide (ANP), its therapeutic use, and processes for producing active ANP.

Background of the Invention

10 Although the production of proteins and peptides by recombinant expression has become relatively common, the expression and subsequent purification of proteins and peptides in an active form having sufficient purity for pharmaceutical administration or diagnostic assay is still a difficult and complicated task. When one is further constrained to use of only commercially feasible processes, the task is perplexing.

15 Recombinant expression of peptides is typically accomplished by inserting a DNA sequence encoding the desired peptide into an expression vector. The expression vector generally contains regulatory sequences which are recognized by the host cell, which provide for transcription and translation of the inserted DNA into the peptide. The expression vector is then inserted into a suitable host cell, typically a bacterium, yeast, or mammalian cell in culture. The expression vector also usually includes a selectable marker, so that one may identify those cells which have been successfully transformed and carry the vector, and separate them from those which do not carry the vector.

20 Peptides may be expressed either "directly" or in the form of a fusion protein. Direct expression produces the desired peptide without modification, but often results in low yields with small peptides. Most host cells appear able to recognize heterologous peptides and effect their degradation.

25 Fusion proteins comprise a leader protein of significant size which is joined to the peptide of interest. The leader protein is often native to the host, as in the case of  $\beta$ -galactosidase. The fusion protein results from expression of a vector containing the coding sequence for the leader protein coupled to the coding sequence for the peptide, joined in the same reading frame. Fusion proteins are often used with small peptides, as the use of the larger endogenous leader protein tends to minimize degradation of the heterologous peptide. However, purification of the fusion protein is often difficult as it often forms insoluble inclusion bodies in the host cell. Use of an endogenous leader protein also hampers separation of the fusion protein from other proteins endogenous to the cell. Further, in order to obtain an authentic peptide one must cleave the leader protein and separate it from the desired peptide. Thus, significantly more processing is required, which leads to loss of product.

30 F.J. Bailey et al, J Indust Microbiol (1987) 2:47-52 disclosed expression of atrial natriuretic peptide (ANP) in *E. coli* as a fusion protein using CheY as the leader protein.

35 Brewer et al, U.S. Pat. No. 4,880,911, disclosed expression of urogastrone (epidermal growth factor) as a fusion protein with a highly-charged amino acid polymer. The peptide is expressed with a C-terminal extension of basic amino acids, preferably 2-30 Arg residues. Following expression, the fusion protein may be isolated using the high positive charge imparted by the poly-Arg tail. The tail may be removed using an exopeptidase such as carboxypeptidase C.

40 W.L. Sung et al, Proc Nat Acad Sci USA (1986) 83:561-65 disclosed expression of proinsulin as fusion proteins having a short (eight amino acid)  $\beta$ -galactosidase leader followed by six or seven repeats of each amino acid, fused to the amino terminus of proinsulin by a linker cleavable with CNBr. Sung found that fusion proteins wherein the hexamer or heptamer was Gln, Asn, Thr, Ser, Ala, His, or Cys (e.g.,  $\beta$ -gal-Gln<sub>6</sub>-proinsulin) were expressed to the greatest degree.

45 Shen et al, EP 163,406, disclosed a method for expressing small peptides as fusion proteins containing a leader protein and multiple copies of the peptide in tandem repeats. Shen used  $\beta$ -galactosidase as the leader protein, and separated multiple copies of insulin with cleavable linker sequences (cleaving with trypsin and carboxypeptidase B). See also S. Cockle et al, "Protein Purification: Micro to Macro" (1987, Alan R. Liss, Inc) pp. 375-81.

50 Cohen et al, U.S. Pat. No. 4,743,679 disclosed the expression of epidermal growth factor (EGF) as a fusion protein with a leader of up to 200 amino acids (preferably up to 75). The leader and EGF peptides are joined by a Glu residue, providing for cleavage by Staph V8. The tertiary structure of EGF apparently protected its internal Glu residues from cleavage by V8. The fusion protein was expressed in bacteria as an insoluble inclusion body. The leader polypeptide was selected to have only one Met (at the N-terminal) and one Glu (where it was joined to EGF) and to exclude Cys.

55 Hobden et al, GB 2,180,539, disclosed a method for producing ANP by expressing the peptide as a fusion protein with a carrier polypeptide. The carrier polypeptide reduced degradation of the short ANP peptide. The fusion protein consisted of a carrier polypeptide joined to the ANP peptide through a linker containing a site cleaved by a protease. The proteases disclosed were mammalian gut proteolytic enzyme enterokinase, thrombin, plasmin, collagenase, Staph V8, Factor Xa and endopeptidase lys C. The preferred carrier polypeptide was derived from the *E. coli*/chloramphenicol acetyltransferase (CAT) gene. Hobden suggested that the fusion protein may be purified prior to proteolysis, and may be assayed based on the immunological characteristics of the carrier polypeptide.

Mai et al, EP 207,044 disclosed the expression of small peptides as fusion proteins comprising an endogenous polypeptide leader, an endopeptidase "trigger signal" (cleavage site), and the peptide. Mai taught that it is preferred to avoid trigger signals in the leader protein so that a clean release of the peptide can be achieved. The maintenance of the integrity of the endogenous protein also simplifies subsequent purification of the peptide.

The proteases disclosed in Mai were trypsin, plasmin, enterokinase, kallikrein, urokinase, tPA, clostripain, chymotrypsin, pepsin, chymosin, collagenase, Russell's viper venom protease, post-proline cleaving enzyme, Staph V8, factor Xa, and thrombin, with Staph V8, factor Xa, and thrombin being preferred. The endogenous proteins suitable for use as leaders disclosed were chloramphenicol acetyltransferase (CAT),  $\beta$ -galactosidase, and recA.

RecA is a nuclear protein found in *E. coli* and other organisms which interacts with DNA, and has a high negative charge. Mai disclosed that using recA as the leader protein improves purification of the subsequent fusion protein, as one can employ techniques such as anion exchange chromatography to isolate recA proteins from other bacterial proteins. Mai separated recA fusion proteins from other bacterial proteins by applying the cell paste extract to an anion exchange column and eluting with either a NaCl gradient or an  $\text{NH}_4\text{Cl}$  gradient, which removed ~80% of the bacterial proteins. After the fusion protein was cleaved with V8 protease, the heterologous protein was separated from the recA portion either by binding to an anion exchange resin or by reversed phase HPLC.

Atrial natriuretic peptide (ANP) is a 3 kDa peptide found in atrial muscle cells which exhibits potent diuretic and natriuretic activities, as well as vasorelaxation. It is normally expressed in the form of a propeptide, having a long pro segment (14-16 kDa) at the amino terminus. ANP is disclosed in Johnson et al, U.S. Pat. No. 4,764,504, filed 8 May 1985, and USSN 870,795, filed 5 June 1986.

#### Disclosure of the Invention

We have now developed an improved method for expressing peptides as fusion proteins, by using the pro portion of human proANP as the carrier for a heterologous peptide, wherein each of the Glu residues normally present in the proANP pro portion is altered to Gln, for example, by site-directed mutagenesis. The proANP pro portion and heterologous peptide are joined by a V8 cleavage site, which permits cleavage using *Staphylococcus aureus* V8 protease. Alternatively, other proteases and chemical cleavage methods can be used to obtain the final peptide of the invention, using the peptide expression methods described herein. The altered proANP pro portion exhibits a high isoelectric point, which facilitates separation of the fusion protein from all other host cell proteins, nucleic acids, pyrogens, and the like, and separation of the carrier from the peptide after cleavage.

The present invention provides a fusion protein which comprises, from N-terminal to C-terminal:

a carrier protein of 10 to 50 kDa which does not contain accessible Glu residues or Asp-Gly sequences as a Staph V8 cleavage site;

a Staph V8 cleavage site comprising a Glu residue or Asp-Gly sequence positioned at the C-terminal of said carrier; and

a peptide not containing Glu residues or Asp-Gly sequences accessible as a cleavage site, fused to said cleavage site, wherein said fusion protein exhibits a pI of 8.0 or greater.

The invention further provides DNA encoding a fusion protein of the invention, and an expression vector for recombinant expression of, in a suitable host, a fusion protein wherein said vector comprises:

a promoter functional in said host;

a DNA of the invention; and

a termination sequence.

The invention also provides a host cell for recombinantly expressing a fusion protein wherein said host cell comprises either a DNA construct comprising:

a promoter functional in said host cell;

a DNA of the invention; and

a termination sequence;

or said host cell comprises a vector of the invention.

The invention yet further provides a method for producing a purified peptide, said method comprising:

expressing a fusion protein within a host cell of the invention;

isolating said fusion protein from proteins endogenous to said host cell by a process which comprises cation ex-

change chromatography under conditions wherein said fusion protein is bound to a cation exchange column while proteins endogenous to said host cells are not bound to said column;  
 eluting said fusion protein;  
 cleaving said carrier protein from said peptide; and  
 separating said peptide from said carrier protein.

#### Brief Description of the Drawings

Figure 1 provides the nucleotide sequence and amino acid sequence for the vector pHNF117. The  $\beta$ -gal-(thr)6 leader peptide amino acid sequence is listed in italics, while the pro-ANP sequence is shown in normal type. The ANP portion of the amino acid sequence is shown in bold type. Residues which have been altered from Glu to Gln are underlined.

Figure 2 lists the nucleotide sequence and amino acid sequence for the vector pHNF120-1. The  $\beta$ -gal-(thr)6 leader peptide amino acid sequence is listed in italics, while the pro-ANP sequence is shown in normal type. The ANP portion of the amino acid sequence is shown in bold type. Residues which have been altered from Glu to Gln are underlined.

Figure 3 shows the synthetic oligonucleotides used to prepare an ANP coding sequence having a Staph V8 cleavage site.

Figure 4 shows the synthetic oligonucleotides used to prepare a  $\beta$ -gal-(thr)6 leader peptide coding sequence having a Staph V8 cleavage site used in pHNF117, as described in Example 1(b).

Figure 5 shows the synthetic oligonucleotides used to prepare a synthetic proANP peptide coding sequence having a Staph V8 cleavage site used in pHNF117, as described in Example 1(c).

Figure 6 shows the synthetic oligonucleotides used to prepare a synthetic proANP peptide coding sequence having a Staph V8 cleavage site used in pHNF117, as described in Example 1(d).

Figure 7 depicts three chromatograms obtained by RP-HPLC. Figure 7(A) shows the chromatogram obtained from an ANP fusion protein of the invention purified on carboxymethyl-sepharose®. Figure 7(B) shows the chromatogram obtained from an ANP fusion protein of the invention obtained by RP-HPLC pre- and post-cleavage with Staph-V8 protease. Figure 7(C) shows the chromatogram of purified ANP obtained from an ANP fusion protein of the invention.

#### Modes of Carrying Out The Invention

##### A. Definitions

The term "fusion protein" as used herein refers to a chimeric polypeptide having a general form:

(leader)-carrier-cleavage site-heterologous peptide, wherein "heterologous peptide" denotes a polypeptide of up to about 10 kDa which exhibits biological activity, antigenic activity, or the like; "cleavage site" indicates an amino acid sequence which is hydrolyzed by *Staphylococcus aureus* V8 protease; "carrier" denotes a polypeptide of 10 to 50 kDa; and "leader" denotes an optional leader sequence. In some cases the leader sequence may be used for directing secretion of the fusion protein.

The term "heterologous peptide" as used herein will generally refer to a peptide which is not endogenous to the host selected, although this definition will also include endogenous peptides in cases in which overexpression of such is desired. Heterologous peptides are short relative to most proteins, generally having a molecular weight of less than about 10 kDa, and may be glycosylated, sialylated, phosphorylated, or the like. The peptide will also exhibit some form of useful activity, typically either biological activity (for example as a peptide hormone), or antigenic activity, for use in recombinant vaccines and/or immunological assays. The peptide will not include an accessible V8 cleavage site, so that the peptide is not fragmented during separation from the carrier protein. The peptide may either omit any cleavage site, or may express a site in an inaccessible portion of the peptide (e.g., at a position of the peptide which is masked by another portion of the peptide, or by glycosylation, phosphorylation, or the like). Peptides which naturally include a cleavage site for the selected protease may be altered, e.g., by site-specific mutagenesis, to a form in which the site is no longer present in cases where the activity of the peptide may be preserved. Representative peptides within the scope of the invention include, without limitation, atrial natriuretic peptide (ANP), brain natriuretic peptide, somatostatin, glucagon-like peptide, calcitonin, lung surfactant, insulin, growth hormone releasing factor (GRF), bradykinins, endorphins, enkephalins, and the like.

The term "carrier protein" refers to a protein which stabilizes expression of a heterologous peptide, and, together with a heterologous peptide forms the fusion protein of the invention. The carrier protein will generally have a molecular weight of 10 to 50 kDa, and preferably will not have an accessible internal cleavage site. The carrier protein also serves to raise the pI of the resulting fusion protein to a range that permits facile separation of the fusion protein from endogenous host cell proteins. A pI of 8.0 or greater is presently preferred. The carrier protein may be derived from an endogenous host protein or fragment thereof, or may consist of a random sequence of high basicity. Preferably, the

carrier protein will not contain Glu or Asp-Gly residues, to avoid cleavage by V8 into a large number of fragments. A presently preferred carrier protein is derived from the pro region of human proANP by altering each Glu residue to a neutral or basic amino acid. The preferred carrier protein has essentially the following sequence:

Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-  
 Leu-Met-Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-  
 X<sub>1</sub>--X<sub>2</sub>--Lys-Met-Pro-Leu-X<sub>3</sub>--Asp-X<sub>4</sub>--Val-Val-  
 Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>--Pro-Asn-X<sub>6</sub>--X<sub>7</sub>--  
 Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>--Val-  
 Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>--Val-Ser-Pro-Ala-Gln-  
 Arg-X<sub>10</sub>-X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-  
 Asp-Ser-Ser-Asp-Arg-Ser-Ala-Leu-Leu-Lys-Ser-

Lys-Leu-Arg-Ala-Leu-Leu-Thr-Ala-Pro-Arg-Ser-  
 Leu-Lys-Phe-;

where X<sub>1-9</sub> are neutral or basic amino acids, and X<sub>10</sub> and X<sub>11</sub> are amino acids other than Glu, selected to avoid the dipeptide sequence Asp-Gly (i.e., X<sub>11</sub> is not Asp, and if X<sub>10</sub> is Asp then X<sub>11</sub> is not Gly or Asp). In the presently-preferred embodiment, X<sub>1-9</sub> are each Gln.

The fusion protein of the invention may further comprise an N-terminal leader sequence of 6 to 20 amino acids, comprising for example 3 to 9 Thr residues. In one embodiment, the leader consists essentially of the sequence:

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr-  
 Gln-Phe-Arg-Met-.

#### B. General Method

##### Vector Construction:

Construction of suitable vectors containing the desired coding and control sequences employs standard ligation and restriction techniques which are well understood in the art. Isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, tailored and religated in the form desired.

Site specific DNA cleavage is performed by treatment with a suitable restriction enzyme (or enzymes) under conditions which are generally understood in the art, generally following the manufacturer's directions. See, e.g., T. Maniatis et al, "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1982). In general, about 1 µg of plasmid or DNA sequence is cleaved by one unit of enzyme in about 20 µL of buffer solution; in the examples herein, typically, an excess of restriction enzyme is used to insure complete digestion of the DNA substrate. Incubation times of about 1 hr to 2 hr at about 37°C are workable, although variations can be tolerated. After each incubation, protein is removed by extraction with phenol/chloroform, and may be followed by diethyl ether extraction, and the nucleic acid recovered from aqueous fractions by ethanol precipitation followed by separation over a Sephadex® G-50 spin column. If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in Meth Enzymol (1980) 65:499-560.

Restriction cleaved fragments may be blunt ended by treating with the large fragment of *E. coli* DNA polymerase I (Klenow) in the presence of the four deoxyribonucleotide triphosphates (dNTPs) using incubation times of about 15 to 25 min at 20 to 25°C in 50 mM Tris, pH 7.6, 50 mM NaCl, 6 mM MgCl<sub>2</sub>, 6 mM DTT and 5-10 µM dNTPs. The Klenow fragment fills in at 5' sticky ends but chews back protruding 3' single strands, even though the four dNTPs are present. If desired, selective repair can be performed by supplying only 1-3 of the dNTPs, within the limitations dictated by the

nature of the sticky ends. After treatment with Klenow fragment, the mixture is extracted with phenol/chloroform and ethanol precipitated followed by running over a Sephadex® G-50 spin column. Treatment under appropriate conditions with S1 nuclease results in hydrolysis of any single-stranded portion.

Synthetic oligonucleotides are prepared by the triester method of Matteucci et al (J Am Chem Soc (1981) 103:3185) or using commercially available automated oligonucleotide synthesizers. Kinasing of single strands prior to annealing or for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 0.1 nmole substrate in the presence of 50 mM Tris, pH 7.6, 10 mM MgCl<sub>2</sub>, 5 mM dithiothreitol, 1-2 mM ATP, 1.7 pmoles <sup>32</sup>P-ATP (2.9 mCi/mmmole), 0.1 mM spermidine, and 0.1 mM EDTA.

Ligations are performed in 15-30 µL volumes under the following standard conditions and temperatures: 20 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 33 µg/mL BSA, 10 mM-50 mM NaCl, and either 40 µM ATP, 0.01-0.02 (Weiss) units T4 DNA ligase at 0°C (for "sticky end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14°C (for "blunt end" ligation). Intermolecular "sticky end" ligations are usually performed at 33-100 µg/mL total DNA concentrations (5-100 nM total end concentration). Intermolecular blunt end ligations (usually employing a 10-30 fold molar excess of linkers) are performed at 1 µM total ends concentration.

In vector construction employing "vector fragments", the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) in order to remove the 5' phosphate and prevent religation of the vector. BAP digestions are conducted at pH 8 in approximately 150 mM Tris, in the presence of Na<sup>+</sup> and Mg<sup>2+</sup> using about 1 unit of BAP per µg of vector at 60°C for about 1 hr. In order to recover the nucleic acid fragments, the preparation is extracted with phenol/chloroform and ethanol precipitated and desalted by application to a Sephadex® G-50 spin column. Alternatively, religation can be prevented in vectors which have been double digested by additional restriction enzyme digestion of the unwanted fragments.

For portions of vectors derived from cDNA or genomic DNA which require sequence modifications, site specific primer directed mutagenesis may be used. This is conducted using a synthetic oligonucleotide primer complementary to a single-stranded phage DNA to be mutagenized except for limited mismatching, representing the desired mutation. Briefly, the synthetic oligonucleotide is used as a primer to direct synthesis of a strand complementary to the phage, and the resulting double-stranded DNA is transformed into a phage-supporting host bacterium. Cultures of the transformed bacteria are plated in top agar, permitting plaque formation from single cells which harbor the phage.

Theoretically, 50% of the new plaques will contain the phage having the mutated form as a single strand; 50% will have the original sequence. The resulting plaques are hybridized with kinased synthetic primer under allele-specific conditions. In general, one may vary the temperature, ionic strength, and concentration of chaotropic agent(s) in the hybridization solution to obtain conditions under which substantially no probes will hybridize in the absence of an "exact match." For hybridization of probes to bound DNA, the empirical formula for calculating optimum temperature under standard conditions (0.9 M NaCl) is

$$T(^{\circ}\text{C}) = 4(N_G + N_C) + 2(N_A + N_T) - 5^{\circ}\text{C},$$

where N<sub>G</sub>, N<sub>C</sub>, N<sub>A</sub>, and N<sub>T</sub> are the numbers of G, C, A, and T bases in the probe (J. Meinkoth et al, Anal Biochem (1984) 138:267-84).

Plaques which hybridize with the probe are then picked, cultured, and the DNA recovered.

#### Verification of Construction:

Correct ligations for plasmid construction may be confirmed by first transforming *E. coli* strain MM294 obtained from *E. coli* Genetic Stock Center, CGSC #6135, or other suitable host with the ligation mixture. Successful transformants are selected by ampicillin, tetracycline or other antibiotic resistance or using other markers depending on the plasmid construction, as is understood in the art. Plasmids from the transformants are then prepared according to the method of D.B. Clewell et al, Proc Nat Acad Sci USA (1969) 62:1159, optionally following chloramphenicol amplification (D.B. Clewell, J Bacteriol (1972) 110:667). The isolated DNA is analyzed by restriction and/or sequenced by the dideoxy method of F. Sanger et al, Proc Nat Acad Sci USA (1977) 74:5463 as further described by Messing et al, Nucleic Acids Res (1981) 9:309, or by the method of Maxam et al, Methods in Enzymol (1980) 65:499.

#### Expression:

The proteins of the invention may be expressed in either prokaryotic or eukaryotic systems. Prokaryotes are most frequently represented by various strains of *E. coli*. However, other microbial strains may also be used, such as bacilli (for example *Bacillus subtilis*), various species of *Pseudomonas*, and other bacterial strains. In such prokaryotic systems, plasmid vectors which contain replication sites and control sequences derived from a species compatible with

the host are used. For example, *E. coli* is typically transformed using derivatives of pBR322, a plasmid derived from an *E. coli* species by Bolivar et al, Gene (1977) 2:95. Commonly used prokaryotic control sequences, which are defined herein to include promoters for transcription initiation, along with ribosome binding site sequences, include such commonly used promoters as the  $\beta$ -lactamase (penicillinase) and lactose (lac) promoter systems (Chang et al, Nature (1977) 198:1056), the tryptophan (trp) promoter system (Goeddel et al, Nuc Acids Res (1980) 8:4057), and the lambda-derived  $P_L$  promoter and N-gene ribosome binding site (Shimatake et al, Nature (1981) 292:128). However, any available promoter system compatible with prokaryotes can be used.

The expression systems useful in eukaryotic systems of the invention comprise promoters derived from appropriate eukaryotic genes. A class of promoters useful in yeast, for example, include promoters for synthesis of glycolytic enzymes, including those for 3-phosphoglycerate kinase (Hitzeman et al, J Biol Chem (1980) 255:2073). Other promoters include those from the enolase gene (M.J. Holland et al, J Biol Chem (1981) 256:1385) or the Leu2 gene obtained from YEp13 (J. Broach et al, Gene (1978) 8:121).

Suitable mammalian promoters include the early and late promoters from SV40 (Fiers et al, Nature (1978) 273:113) or other viral promoters such as those derived from polyoma, adenovirus II, bovine papilloma virus, or avian sarcoma viruses. Suitable viral and mammalian enhancers are cited above. In the event plant cells are used as an expression system, the nopaline synthesis promoter is appropriate (A. Depicker et al, J Mol Appl Gen (1982) 1:561). Expression in insect cell culture may conveniently be achieved using a baculovirus vector, for example, transfer vectors derived from the baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV), (see PCT W089/046699).

#### Transformation:

Depending on the host cell used, transformation is accomplished using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by S.N. Cohen, Proc Nat Acad Sci USA (1972) 69:2110, or the RbCl method described in Maniatis et al, *supra*, p. 254, is used for prokaryotes or other cells which contain substantial cell wall barriers. Infection with *Agrobacterium tumefaciens* (C.H. Shaw et al, Gene (1983) 23:315) is used for certain plant cells. For mammalian cells without cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology (1978) 52:546 is preferred. Transformations into yeast are carried out according to the method of P. Van Solingen et al, J Bacter (1977) 130:946 and C.L. Hsiao et al, Proc Nat Acad Sci USA (1979) 76:3829.

#### C. Examples

The examples presented below are provided as a further guide to the practitioner of ordinary skill in the art, and are not to be construed as limiting the invention in any way.

#### Example 1

##### (Construction of plasmid pHNF75)

A plasmid was constructed which fuses human atrial natriuretic peptide (hANP4-28), to *E. coli*  $\beta$ -galactosidase (amino acid 1005 at the EcoRI site) with an Endoproteinase Glu-C (Staph V8) cleavage site (Boehringer Mannheim) at their junction. The hybrid gene is transcribed from the *lacZ* promoter/operator region on a pBR322-based plasmid.

First, a synthetic DNA fragment encoding the cleavage site, hANP(4-28), and translation termination signal was assembled from eight oligodeoxyribonucleotides (Figure 3) by the method of G.P. Vlasuk et al, J Biol Chem (1986) 261:4789-96, and purified from a 8% polyacrylamide gel (T. Maniatis et al, *supra*, p. 173-178). Some *E. coli* preferred codon choices (M. Gouy et al, Nuc Acids Res (1982) 10:7055-74) were utilized in the design of the synthetic gene which was designated hANP\* (the asterisk denoting the presence of the proteolytic cleavage site). About 200 ng of this fragment was mixed with 20 ng of M13mp19 (C. Yanisch-Perron et al, Gene (1985) 33:103-19) which had been digested with restriction endonucleases EcoRI and SmaI (New England Biolabs, using conditions recommended by manufacturer) to linearize the vector; 400 U of T4 DNA ligase (New England Biolabs) and ligase buffer (final concentration 50 mM Tris-HCl pH 7.4, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 0.3 mM ATP) were added and the DNAs incubated overnight at 15°C. *E. coli* JM101 (0.1 mL) made competent by the CaCl<sub>2</sub> procedure (Maniatis et al, *supra*, p. 250-51) was transfected with the ligation, mixed with 0.2 mL of JM101 overnight culture, added to L-top agarose with 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside, i.e., X-gal (50  $\mu$ L of 2% solution in dimethylformamide) and isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) (10  $\mu$ L of 100 mM solution in water) and poured on the surface of an L-plate. After overnight incubation at 37°C, a white plaque was chosen (M13-hNF7), a phage stock grown, and single-stranded DNA prepared (J. Messing, Meth Enzymol (1983) 101:20-78) for sequencing by the method of F. Sanger et al, Proc Nat Acad Sci USA (1977) 74:5463-67. The DNA, sequence of the synthetic hANP\* fragment was shown to be correct.

Second, double-stranded DNA from M13-hNF7 (Messing, *supra*) was prepared and digested with restriction endonucleases BamHI and BglII (New England Biolabs) and the approximately 700 bp DNA fragment containing the hANP\* sequence was purified from a 0.6% agarose gel (Maniatis et al, *supra*, pp. 157-161). Plasmid pTrp-233 (prepared as described in US 4,764,504) was digested with restriction endonuclease BamHI, its termini dephosphorylated by incubation with bacterial alkaline phosphatase (Amersham) by the method of Maniatis et al (*supra*, p. 133), and the linearized plasmid purified from a 0.6% agarose gel. Approximately 100 ng of the BamHI-BglII DNA fragment was mixed with 20 ng of the linearized plasmid and 400 U of T4 DNA ligase and ligation buffer added for overnight incubation at 15°C. Competent *E. coli* MC1061 was transformed with the ligation and incubated overnight on L-plates with 100 µg/mL ampicillin at 37°C. Colonies were inoculated into L-broth containing 100 µg/mL ampicillin, grown overnight, and 1 mL withdrawn to prepare plasmid DNA by the alkaline lysis method of Maniatis et al, (*supra*, p. 368-369). The orientation of the BamHI-BglII fragment within the plasmid was determined by digestion with restriction endonucleases PvuII and HindIII (New England Biolabs) and sizing of the DNA fragments on a 0.6% agarose gel. The desired orientation placed the BamHI recognition site of the fragment next to the EcoRI recognition site of the plasmid creating an hANP\* cassette which could be released as an EcoRI-HindIII, EcoRI-BamHI, or EcoRI-EcoRI fragment. A plasmid with this orientation was designated pHNF73.

Next, pHNF73 was digested with restriction endonuclease EcoRI and the approximately 100 bp EcoRI-EcoRI fragment encoding hANP\* was purified from an 8% polyacrylamide gel. Next 2 µg of plasmid pBgal was digested with restriction endonuclease EcoRI, the linearized plasmid purified from a 0.6% agarose gel, and its termini dephosphorylated by incubation with bacterial alkaline phosphatase. Then about 20 ng of linearized plasmid was mixed with about 50 ng of the EcoRI-EcoRI fragment and 400 U of T4 DNA ligase and ligation buffer added for overnight incubation at 15°C. Competent *E. coli* MC1061 were transformed with the ligation and incubated overnight at 37°C on L-plates with 100 µg/mL ampicillin. One plasmid shown to contain the EcoRI-EcoRI fragment by digestion with EcoRI was designated pHNF75. In several *E. coli* strains, pHNF75 was shown to produce a hybrid protein of the expected size for a fusion of β-galactosidase and hANP(4-28) using SDS polyacrylamide gel electrophoresis (U. Laemmli, *Nature* (1970) 227: 680-85). The fusion protein was partially purified, digested with Endoproteinase Glu-C (Staph V8), purified, and shown to have the correct amino acid sequence (data not shown).

#### (b) Construction of pTRP83-1:

Plasmid pTRP83-1, which expresses a short β-gal leader peptide from the tryptophan promoter was constructed as follows:

Plasmid pTrp-233 (prepared as described in US 4,764,504) was used to construct an expression vector encoding five amino-terminal residues of *E. coli* β-galactosidase followed by the sequence -Asn-Leu-Thr-Thr-Thr-Thr-Thr-Glu-Phe- (i.e., β-gal-(thr)6 leader peptide) under the control of the *E. coli* tryptophan promoter/operator region on a pBR322-based plasmid. Figure 4 illustrates preparation of this sequence from four synthetic oligonucleotides. If desired, one may substitute the sequence -Asn-Leu-Thr-Thr-Thr-Thr-Gln-Phe- for the above by altering the appropriate oligonucleotides shown in Figure 4. EcoRI and HindIII recognition sites downstream of this sequence permit fusion of genes, in frame, with the short leader peptide. This and other related peptides have been shown to stabilize the expression of proinsulin in *E. coli* (W.L. Sung et al, *Proc Nat Acad Sci USA* (1986) 83:561-65).

First, 2 µg of pTrp-233 were digested with EcoRI and the linearized plasmid purified from a 0.6% agarose gel. The termini were then filled in by the method of Maniatis et al (*supra*, p. 394), using the Klenow fragment of *E. coli* DNA polymerase I (Boehringer-Mannheim, Inc.). The mixture was then heated to 70°C for 5 min to inactivate the enzyme and T4 DNA ligase was added with ligation buffer for overnight incubation at 15°C. *E. coli* MC1061 were made competent by the CaCl<sub>2</sub> method and transformed as described by Maniatis et al (*supra*, pp. 250-251). Resulting ampicillin-resistant colonies were grown overnight in L-broth with 100 µg/mL ampicillin and 1 mL aliquots withdrawn to prepare plasmid DNA by the alkaline lysis method of Maniatis et al (*supra*, pp. 368-369). A plasmid which had lost the EcoRI recognition site, as determined by failure to be digested by EcoRI, was designated pTRP81-6.

Next, 2 µg of plasmid pTRP81-6 was digested with HindIII and NdeI and linearized plasmid purified from a 0.6% agarose gel. A synthetic NdeI-HindIII DNA fragment encoding the β-gal-(thr)6 leader peptide was assembled from four oligodeoxyribonucleotides (Figure 4) using the method of Vlasuk et al (*supra*) and purified from a 20% polyacrylamide gel (Maniatis et al, *supra*). This fragment encodes EcoRI and HindIII recognition sites for the fusion of genes, in frame, with β-galactosidase and the termini are nonphosphorylated. A thousandfold excess of the synthetic fragment was added to the linearized plasmid, mixed with T4 DNA ligase and ligation buffer, and incubated overnight at 15°C; the ligation was used to transform competent *E. coli* MC1061 to ampicillin resistance. Colonies were grown, plasmid DNA prepared, and an EcoRI digest used to identify a plasmid containing the synthetic fragment, designated pTRP83-1. The correct sequence of the synthetic DNA fragment encoding the leader peptide was confirmed by the method of Sanger et al (*supra*).



(c) Construction of pHNF86:

Plasmid pHNF86, which expresses a  $\beta$ -gal-(thr)6-proANP\* hybrid protein from the tryptophan promoter, was constructed as follows:

Plasmids pHNF-233 (prepared as in US 4,764,504), pHNF75 (see section a above), and pTRP83-1 (see section b above) were used to construct an expression vector encoding the  $\beta$ -gal-(thr)6 leader peptide followed by proANP8 which contains an Endoproteinase Glu-C (Staph V8) cleavage site between pro and hANP(4-28) sequences. The hybrid protein is under the control of the *E. coli* tryptophan promoter/operator region on a pBR322-based plasmid.

First, plasmid pHNF-233 was partially digested with Aval and completely digested with HindIII to linearize the plasmid and remove the 3'-end of the proANP gene (from Aval to HindIII). The linearized plasmid was purified from a 0.6% agarose gel. Plasmid pHNF75 was digested with EcoRI and HindIII and the fragment encoding the hANP(4-28) gene with Endoproteinase Glu-C (Staph V8) cleavage site was purified from a 20% polyacrylamide gel. A synthetic DNA fragment encoding proANP sequences from the Aval recognition site to the EcoRI site of the hANP\* fragment (Figure 5) was assembled from eight oligodeoxyribonucleotides by the method of Vlasuk et al (*supra*) and purified from an 8% polyacrylamide gel. The DNA sequence had been modified on this fragment so that ligation of the single-stranded tail with the complementary EcoRI tail destroys the EcoRI site (GAATTC changed to AAATTC); the termini are also nonphosphorylated. The three DNAs were mixed (about 50 ng of plasmid backbone, 100 ng of synthetic fragment, and at least 100 ng of hANP\* fragment) with T4 DNA ligase and ligation buffer, and incubated overnight at 15°C. The ligation used to transform competent *E. coli* MC1061, ampicillin-resistant colonies grown in culture and plasmid DNA prepared for restriction endonuclease digestion with EcoRI and HindIII to confirm the insertion of the two DNA fragments into the plasmid. The correct DNA sequence was confirmed for the new proANP\* using the method of Sanger et al (*supra*) in a plasmid designated pHNF82.

Next, plasmid pHNF82 was digested with NdeI and the termini filled in with the Klenow fragment of *E. coli* DNA polymerase I; the blunted linearized plasmid was mixed with a nonphosphorylated EcoRI linker (5'-CGGAATTCCG-3', New England Biolabs) and T4 DNA ligase for overnight incubation at 15°C. The ligation was then extracted with phenol: chloroform, ethanol precipitated, and digested with EcoRI and HindIII to release the proANP gene as an EcoRI-HindIII fragment purified from a 0.6% agarose gel. Plasmid pTRP83-1 was digested with EcoRI and HindIII to linearize the plasmid and purified from a 0.6% agarose gel. The DNA fragment and the linearized plasmid were mixed with T4 DNA ligase and ligation buffer and incubated overnight at 15°C. Competent *E. coli* MC1061 were transformed with the ligation and ampicillin-resistant colonies grown as overnight cultures for plasmid DNA. Digestion of plasmid DNA with EcoRI and HindIII identified a plasmid containing the proANP gene, designated pHNF86. Upon the addition of 3- $\beta$ -indoleacrylic acid (25  $\mu$ g/mL) to induce the tryptophan promoter, several strains of *E. coli* carrying plasmid pHNF86 expressed the hybrid protein at 10-15% of total cell protein in shake flask cultures (data not shown). The expression level was estimated by densitometer scan of Coomassie blue-stained SDS polyacrylamide gels.

(d) Construction of plasmid pHNF117:

Plasmid pHNF117, which expresses a  $\beta$ -gal-(thr)6-proANP\* gln hybrid protein from the tryptophan promoter, was constructed as follows:

Plasmid pHNF86 was modified to reduce the number of endoproteinase Glu-C (Staph V8) cleavage sites (Boehringer Mannheim) within the proANP region. Only two sites remain in the modified plasmid: the first at the junction between the  $\beta$ -gal-(thr)6 leader peptide and the proANP region and the second between the proANP region and the hANP(4-28) peptide (Figure 1). This was accomplished by changing nine glutamic acid residues to glutamine residues; preferred *E. coli* codon choice (Gouy et al, *supra*) was also introduced in the region which was modified.

First, plasmid pHNF86 was digested with EcoRI and HindIII and the fragment encoding proANP\* purified from a 0.6% agarose gel. Plasmid pUC8 (J. Vieira et al, *Gene* (1982) 19:252) was digested with EcoRI and HindIII and the linearized plasmid purified from a 0.6% agarose gel. DNA fragment (500 ng) and linearized plasmid (100 ng) were mixed, incubated with T4 DNA ligase and ligation buffer overnight at 15°C and the ligation used to transform competent *E. coli* MC1061 to ampicillin resistance. A plasmid containing the proANP\* fragment was designated pUC-hNF114.

Next, the EcoRI-Aval DNA sequence of the proANP fragment was removed by digestion with EcoRI and Aval and purification of the plasmid on a 0.6% polyacrylamide gel. A synthetic DNA fragment to replace this region was assembled from sixteen oligodeoxyribonucleotides using the method of Vlasuk et al (*supra*) (Figure 6). The sequence of the fragment incorporates both preferred *E. coli* codons and the replacement of nine glutamic acid residues by glutamine residues. The nonphosphorylated synthetic DNA fragment was purified from an 8% polyacrylamide gel and at least 200 ng mixed with 50 ng of the linearized plasmid pUC-hNF114. T4 DNA ligase and ligation buffer were added, and the mixture incubated overnight at 15°C. The ligation was used to transform *E. coli* MC1061 to ampicillin resistance. One plasmid, designated pUC-hNF116, was shown to contain the substituted region, proANP\*gln, by the DNA sequencing method of Sanger et al (*supra*).

Plasmid pHNF86 digested with EcoRI and HindIII was then purified as a linear molecule from a 0.6% agarose gel and 50 ng mixed with at least 100 ng of EcoRI-HindIII proANP\**gln* fragment purified on a 0.6% agarose gel from an EcoRI and HindIII digest. Following overnight incubation with T4 DNA ligase and ligation buffer at 15°C, competent *E. coli* MC1061 were transformed to ampicillin resistance, grown overnight in culture, and plasmid DNA prepared. One expression vector, designated pHNF117, encoded the new proANP\**gln* gene fused to the  $\beta$ -gal-(thr)6 leader peptide. *E. coli* W3110 carrying this plasmid was shown to express the hybrid protein in shake flask cultures, upon the addition of 3- $\beta$ -indoleacrylic acid; densitometer tracings of Coomassie blue-stained SDS polyacrylamide gels estimated the hybrid protein at 10% of total cell protein.

#### (e) Construction of plasmid pHNF120-1:

Plasmid pHNF120-1 is identical to pHNF117, except that the Glu appearing in the  $\beta$ -gal-(thr)6 leader is altered to Gln, thus leaving only one Staph V8 cleavage site in the expression product. The plasmid was constructed site-directed mutagenesis of pHNF117, and the product confirmed to have the sequence shown in Figure 2.

#### (f) Construction of plasmid pHNF120-3:

Plasmid pHNF120-3 is identical to pHNF120-1, except that the expression product does not contain any Staph V8 cleavage sites.

#### Example 2

(Purification of recombinant human ANP-fusion protein)

#### Fermentation:

*E. coli* W3110 cells transformed with pHNF117 were fermented by a fed-batch fermentation process in complex media containing an initial concentration of 5.0 g/L glucose. A concentrated glucose feed was started when the residual sugar was measured to be less than 0.5 g/L. The feed rate was then adjusted to maintain a residual glucose concentration of less than 1.0 g/L and a minimum dissolved oxygen of 20%. Fusion protein production was induced by the addition of 100 mg/L indoleacrylic acid (IAA) when the cell density was 40.0 OD<sub>590</sub>. The fermentation was continued for an additional 14.5 h and harvested at a cell density of 73.0 OD<sub>590</sub>. The total fermentation time was about 28.5 h.

#### Cell Recovery:

Eighteen liters of fermentation broth were divided into six 1 liter bottles and centrifuged at 5,000 rpm for 30 min at 4°C in a Sorvall RC-3B centrifuge. The supernatant was discarded, and the bottles refilled with whole fermentation broth and centrifuged again. This process was repeated three times, resulting in a recovery of 1.23 Kg cells (wet weight). All cells were frozen at -85°C until further processing. We have also used a cross-flow microfiltration unit equipped with 0.1  $\mu$ m membranes (Sartorius, Inc., Yauco, P.R.) to reduce the fermentation broth volume so that only one centrifugation step is required to recover the cells.

#### Homogenization:

Two hundred grams of cells were thawed to 4°C and suspended in 2 liters (1:10, w/v) of 0.01 M MES (2-[N-morpholino]ethanesulfonic acid), pH 6.0, for 1 hr at 4°C. The cell suspension was recirculated 3 times at 300 mL/min, with cooling, through a Microfluidizer (Model 110Y, Microfluidics corp., Newton, MA) operated at 10,000 psi. The fluidizer was equipped with 75- and 150-micron mixing chamber to break open the cells and release the intracellular proteins. Cell lysis was judged to be greater than 90% complete by microscopic examination.

The lysate was divided between 1 liter bottles and centrifuged at 5,000 rpm in a Sorvall RC-3B centrifuge for 30 min at 4°C. The pellet containing the insoluble ANP-fusion protein (inclusion bodies) was saved and the supernatant containing the soluble *E. coli* proteins was discarded.

#### Wash Inclusion Bodies:

The lysate pellet (inclusion bodies) was resuspended in 750 mL (1:10, w/v) of 0.5 M urea in 0.04 M H<sub>2</sub>SO<sub>4</sub> with magnetic mixing at 4°C for 30 min. The suspended inclusion bodies were recirculated through the microfluidizer twice, to produce a homogeneous mixture and to break open any residual cells that may have copelleted with the inclusion

bodies. The mixture was centrifuged at 5,000 rpm in a Sorvall RC-3B centrifuge at 4°C for 20 min and the pellet saved. The washed pellet may be stored at -70°C until further processing.

These inclusion bodies were washed with the low pH solution for two reasons: (1) to remove residual soluble proteins, and (2) to inhibit endogenous proteolytic enzymes that have been shown to digest ANP fusion protein at higher pH.

#### Solubilize Inclusion Bodies:

The washed inclusion bodies were suspended in 0.5 L (1:5, w/v) of 6 M urea in 0.4 M H<sub>2</sub>SO<sub>4</sub> with magnetic stirring at 4°C for 30 min and then recirculated through the Microfluidizer twice to homogenize particles and solubilizes the ANP fusion protein. The resulting extract was centrifuged at 5,000 rpm in a Sorvall RC-3B rotor at 4°C for 1 hr to remove residual insoluble materials. The supernatant may be processed immediately or stored at -70°C to await further processing. Alternatively, the extract can be clarified by cross-flow filtration over a 0.2 µm filter, such as the Minitan unit (Millipore, Inc., Bedford, MA) with 500 cm<sup>2</sup> surface area.

#### Cation Exchange Chromatography:

The new ANP fusion protein was shown to have a basic isoelectric point; therefore, cation exchange chromatography was selected to further purify the protein to near homogeneity before Staph V8 cleavage. A sample (0.5 L; 16.0 mg protein/mL) of solubilized ANP fusion protein was diluted to 4.0 L with 6 M urea in 0.01 M MES and the pH was adjusted to 6.0 with sodium hydroxide at 4°C. The diluted sample was loaded at 50 mL/min (38.2 cm/h linear velocity) onto a 10 × 16.5 cm (1.3 L) column of CM (carboxymethyl)-Sephacrose® Fast-Flow (Pharmacia, Inc., Piscataway, NJ) equilibrated with 6 M urea in 0.01 M MES, pH 6.0 at 4°C. (Alternatively, one may employ sulfopropyl-Sephacrose® or other cation exchange materials.) After loading, unbound proteins were washed into the flow-through fraction with additional equilibration buffer until the absorbance at 280 nm returned to baseline. Fusion protein was then eluted with a solution of 6 M urea, a linear gradient from 0 to 0.06 M NaCl in 0.01 M MES (or 0.06 M NaCl in 0.01 M MES), pH 6.0, at 100 mL/min (76.4 cm/hr linear velocity) and collected in 400 mL fractions. Fractions were analyzed for purity by SDS-PAGE before pooling for cleavage, and 2.7 g were recovered (34% step yield). It may be advantageous to use other pH values. A pH of 6.0 was used herein to minimize any carbamylation reactions that may result from urea decomposition at elevated pH.

#### Cleavage with Staph V8:

CM-Sephacrose® purified fusion protein was shown to contain trace levels of *E. coli* proteolytic enzymes which led to product loss and nonspecific fragmentation of the fusion during prolonged incubation with immobilized Staph V8. These nonspecific proteases were inactivated by heating the CM-purified fusion protein to 80°C as follows. First, 0.1 M Tris SO<sub>4</sub>, pH 9.0, is heated to 80°C and mixed with an equal volume of purified fusion protein solution. The solution is maintained at 80°C for 30 min, then cooled to room temperature. During this treatment, the urea concentration is reduced to 3 M so that it will not inhibit the cleavage reaction. If necessary, the volume of this solution may be reduced, prior to cleavage, by ultrafiltration with 10,000 MWCO membranes.

Staph V8 (endoproteinase Glu-C, Boehringer Mannheim GmbH, West Germany) was immobilized onto glutaraldehyde-activated PVC-silica composite membranes (Amerace, Inc., Hackettstown, NJ) to a density of approximately 0.2 mg enzyme/cm<sup>2</sup> membrane. (One may alternatively employ cyanogen bromide-activated Sepharose®, or other activated resins, membranes, or substrates.) Three 47 mm diameter membranes containing approximately 15 mg Staph V8 were connected in parallel and equilibrated with 0.05 M Tris-SO<sub>4</sub>, pH 9.0, at 2 mL/min at room temperature. Heat-treated fusion protein solution (200 mL) containing 1000 mg of protein was recirculated through the membrane system at 2 mL/min for 16 hr at room temperature. Cleavage was monitored by SDS-PAGE and RP-HPLC. Figure 7A shows the RP-HPLC chromatogram of a sample prepared as described above, after purification on carboxymethyl Sepharose®. Figure 7B shows the RP-HPLC chromatogram of the sample after cleavage with Staph V8. One may alternatively effect cleavage using soluble (not immobilized) enzyme.

#### Ultrafiltration Size Separation:

After cleavage with immobilized Staph V8, ANP was separated from the high molecular weight cleavage products by simple size separation. Ultrafiltration of the cleavage mixture over an Amicon YM10 membrane (10,000 MWCO) with washing, provided about 80% recovery of the product in the filtrate stream, free of high molecular weight fragments.

In addition to ANP, any ultrafiltration filtrate stream may contain the N-terminal peptide (when pHNF117 is used), and urea from the cleavage reaction. The N-terminal peptide may be eliminated by replacing the β-gal-(thr)6 leader

peptide with the Glu-Gln variant described above. To separate ANP from these substances, a second cation exchange chromatography step may be used. Alternatively, gel filtration chromatography can be used to separate ANP from high molecular weight cleavage products.

#### 5 Cation Exchange Chromatography:

Cation exchange chromatography was used to separate ANP from its cleavage products after cleavage or after ultrafiltration. In this step, ANP was chromatographically resolved on a CM-Sepharose® Fast Flow column with a salt and pH gradient of ammonium acetate.

10 Alternatively, ANP has been separated by reversed-phase high performance liquid chromatography and lyophilized to remove organic solvents.

#### 15 Claims

Claims for the following Contracting States : AT, BE, CH, DE, DK, FR, GB, GR, IT, LI, LU, NL, SE

1. A fusion protein which comprises, from N-terminal to C-terminal:

20 a carrier protein of 10 to 50 kDa which does not contain accessible Glu residues or Asp-Gly sequences as a Staph V8 cleavage site;

a Staph V8 cleavage site comprising a Glu residue or Asp-Gly sequence positioned at the C-terminal of said carrier; and

25 a peptide not containing Glu residues or Asp-Gly sequences accessible as a cleavage site, fused to said cleavage site, wherein said fusion protein exhibits a pI of 8.0 or greater.

2. The fusion protein of claim 1, wherein said carrier protein consists essentially of the sequence:

30 Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
35 Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
40 X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

45 wherein

X<sub>1-9</sub> are each independently selected from the group consisting of Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, and Tyr; and

50 X<sub>10</sub> and X<sub>11</sub> are each independently selected from the group consisting of Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, and Tyr, with the proviso that X<sub>11</sub> is not Asp, and if X<sub>10</sub> is Asp then X<sub>11</sub> is not Gly.

3. The fusion protein of claim 2, wherein X<sub>1-9</sub> are each Gln.

4. The fusion protein of any one of claims 1 to 3 wherein said fusion protein further comprises an N-terminal leader of 6 to 20 amino acids.

5. The fusion protein of claim 4, wherein said leader comprises from 3 to 9 Thr residues.
6. The fusion protein of claim 5, wherein said leader consists essentially of the sequence:

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr-Gln-Phe-Arg-Met-.

7. The fusion protein of claim 6, wherein said peptide is atrial natriuretic peptide (ANP), an analog of ANP, brain natriuretic peptide, somatostatin, glucagon-like peptide, calcitonin, lung surfactant, insulin, growth hormone releasing factor, a bradykinin, an endorphin or an enkephalin.
8. The fusion protein of any one of claims 1 to 7 wherein said peptide consists essentially of the sequence:

Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-Gly-Ala-Gln-Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-Tyr.

9. A DNA encoding a fusion protein according to any one of claims 1 to 8.

10. An expression vector for recombinant expression of, in a suitable host, a fusion protein wherein said vector comprises:

a promoter functional in said host;  
a DNA according to claim 9; and  
a termination sequence.

11. The vector of claim 10, wherein said promoter comprises the *E. coli*  $\beta$ -galactosidase promoter.

12. The vector of claim 10 or 11 which further comprises a selectable marker.

13. The vector of any one of claims 10 to 12 wherein said vector is a plasmid.

14. The vector of claim 13, which comprises pHNF120-1, as defined in Example 1, section (e).

15. A host cell for recombinantly expressing a fusion protein wherein said host cell comprises either a DNA construct comprising:

a promoter functional in said host cell;  
a DNA according to claim 9; and  
a termination sequence;

or said host cell comprises a vector according to any one of claims 10 to 14.

16. The host cell of claim 15, wherein said DNA or vector is present on an extrachromosomal element.

17. The host cell of claim 15, wherein said DNA or vector is integrated within the genome of said host cell.

18. A method for producing a purified peptide, said method comprising:

expressing a fusion protein within a host cell according to any one of claims 15 to 17;  
isolating said fusion protein from proteins endogenous to said host cell by a process which comprises cation exchange chromatography under conditions wherein said fusion protein is bound to a cation exchange column, while proteins endogenous to said host cells are not bound to said column;

eluting said fusion protein;  
cleaving said carrier protein from said peptide; and  
separating said peptide from said carrier protein.

Claims for the following Contracting State : ES

1. A method of producing a fusion protein which comprises from N-terminal to C-terminal:

a carrier protein of 10 to 50 kDa which does not contain accessible Glu residues or Asp-Gly sequences as a Staph V8 cleavage site;  
a Staph V8 cleavage site comprising a Glu residue or Asp-Gly sequence positioned at the C-terminal of said carrier; and  
a peptide not containing Glu residues or Asp-Gly sequences accessible as a cleavage site, fused to said cleavage site, wherein said fusion protein exhibits a pI of 8.0 or greater; which method comprises:

expressing the fusion protein within a host cell;  
isolating said fusion protein from proteins endogenous said host cell by a process which comprises cation exchange chromatography under conditions wherein said fusion protein is bound to a cation exchange column, while proteins endogenous said host cells are not bound to said column; and  
eluting said fusion protein.

2. A method according to claim 1, wherein said carrier protein consists essentially of the sequence:

Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

wherein

X<sub>1-9</sub> are each independently selected from the group consisting of Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, and Tyr; and  
X<sub>10</sub> and X<sub>11</sub> are each independently selected from the group consisting of Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, and Tyr, with the proviso that X<sub>11</sub> is not Asp, and if X<sub>10</sub> is Asp then X<sub>11</sub> is not Gly.

3. A method according to claim 2, wherein X<sub>1-9</sub> are each Gln.  
4. A method according to any one of claims 1 to 3 wherein said fusion protein comprises an N-terminal leader of 6 to 20 amino acids.  
5. A method according to claim 4, wherein said leader comprises from 3 to 9 Thr residues.  
6. A method according to claim 5, wherein said leader consists essentially of the sequence:

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr-  
Thr-Gln-Phe-Arg-Met-.

7. A method according to claim 6, wherein said peptide is atrial natriuretic peptide (ANP), an analog of ANP, brain natriuretic peptide, somatostatin, glucagon-like peptide, calcitonin, lung surfactant, insulin, growth hormone releasing factor, a bradykinin, an endorphin or an enkephalin.

8. A method according to any one of claims 1 to 7 wherein said peptide consists essentially of the sequence:

Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-  
Gly-Ala-Gln-Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-  
Tyr.

9. A method for the preparation of an expression vector for recombinant expression of, in a suitable host, a fusion protein wherein said vector comprises:

a promoter functional in said host;  
a DNA sequence encoding a fusion protein as defined in any one of claims 1 to 8; and  
a termination sequence;

which method comprises operably linking said promoter, said DNA coding sequence and said termination sequence in a vector capable of directing expression of the DNA sequence in a host cell.

10. A method according to claim 9, wherein said promoter comprises the *E. coli*  $\beta$ -galactosidase promoter.

11. A method according to claim 9 or 10 wherein said vector further comprises a selectable marker.

12. A method according to any one of claims 9 to 11 wherein said vector is a plasmid.

13. A method according to claim 12 wherein said vector comprises pHNF120-1, as defined in Example 1, section (e).

14. A method for the preparation of a host cell for recombinantly expressing a fusion protein which method comprises transforming a host cell with either a DNA construct comprising:

a promoter functional in said host cell;  
a DNA sequence encoding a fusion protein as defined in any one of claims 1 to 8; and  
a termination sequence;

or with a vector according to any one of claims 9 to 13.

15. A method according to claim 14, wherein said DNA sequence or vector is present on an extrachromosomal element.

16. A method according to claim 14, wherein said DNA sequence or vector is integrated within the genome of said host cell.

17. A method for producing a purified peptide, said method comprising producing a fusion peptide comprising a carrier protein, a Staph V8 cleavage site and a peptide according to the method of any one of claims 1 to 8;

cleaving said carrier protein from said peptide; and  
separating said peptide from said carrier protein.

18. A fusion protein which comprises, from N-terminal to C-terminal:

a carrier protein of 10 to 50 kDa which does not contain accessible Glu residues or Asp-Gly sequences as a

Staph V8 cleavage site;

a Staph V8 cleavage site comprising a Glu residue or Asp-Gly sequence positioned at the C-terminal of said carrier; and

a peptide not containing Glu residues or Asp-Gly sequences accessible as a cleavage site, fused to said cleavage site, wherein said fusion protein exhibits a pI of 8.0 or greater.

19. A DNA encoding a fusion protein according to claim 18.

20. An expression vector for recombinant expression of, in a suitable host, a fusion protein wherein said vector comprises:

a promoter functional in said host;

a DNA according to claim 19; and

a termination sequence.

21. A host cell for recombinantly expressing a fusion protein wherein said host cell comprises either a DNA construct comprising:

a promoter functional in said host cell;

a DNA according to claim 19; and

a termination sequence;

or said host cell comprises a vector according to claim 20.

## Patentansprüche

Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, DE, DK, FR, GB, GR, IT, LI, LU, NL, SE

1. Fusionsprotein, vom N-terminalen zum C-terminalen Ende enthaltend:

ein Trägerprotein von 10 bis 50 kDa, das keine zugänglichen Glu-Reste oder Asp-Gly-Sequenzen als Staph V8-Spaltungsstelle enthält;

eine Staph V8-Spaltungsstelle, die einen Glu-Rest oder eine Asp-Gly-Sequenz aufweist, die an dem C-terminalen Ende des Trägers positioniert ist; und

ein Peptid, das keine als Spaltungsstelle zugänglichen Glu-Reste oder Asp-Gly-Sequenzen enthält, wobei das Peptid an die Spaltungsstelle fusioniert ist und das Fusionsprotein einen pI von 8,0 oder größer zeigt.

2. Fusionsprotein gemäß Anspruch 1, wobei das Trägerprotein im wesentlichen aus der Sequenz besteht:



Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
 Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
 5 Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
 Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
 Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
 10 X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
 Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
 Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

worin

X<sub>1-9</sub> jeweils unabhängig aus der aus Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, und Tyr bestehenden Gruppe ausgewählt ist; und

X<sub>10</sub> und X<sub>11</sub> jeweils unabhängig aus der aus Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp und Tyr bestehenden Gruppe ausgewählt sind, unter der Voraussetzung, daß X<sub>11</sub> nicht Asp und, wenn X<sub>10</sub> Asp ist, daß X<sub>11</sub> dann nicht Gly ist.

3. Fusionsprotein gemäß Anspruch 2, wobei X<sub>1-9</sub> jeweils Gln ist.
4. Fusionsprotein gemäß einem der Ansprüche 1 bis 3, wobei das Fusionsprotein ferner am N-terminalen Ende eine Leitsequenz von 6 bis 20 Aminosäuren aufweist.
5. Fusionsprotein gemäß Anspruch 4, wobei die Leitsequenz 3 bis 9 Thr-Reste aufweist.
6. Fusionsprotein gemäß Anspruch 5, wobei die Leitsequenz im wesentlichen aus der Sequenz besteht:

**Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr-Thr-  
 Gln-Phe-Arg-Met-.**

7. Fusionsprotein gemäß Anspruch 6, wobei das Peptid ein atrio-natriuretisches Polypeptid (ANP) ist, ein analoges von ANP, ein Gehirn-natriuretisches Polypeptid, Somatostatin, ein glucagon-ähnliches Peptid, Calcitonin, ein Lungentensid, Insulin, ein Wachstumshormon-Freisetzungsfaktor, ein Bradykinin, ein Endorphin oder ein Enkephalin.
8. Fusionsprotein gemäß einem der Ansprüche 1 bis 7, wobei das Peptid im wesentlichen aus der Sequenz besteht:

**Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-Gly-  
 Ala-Gln-Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-Tyr.**

9. DNA, die ein Fusionsprotein gemäß einem der Ansprüche 1 bis 8 codiert.
10. Expressionsvektor zur rekombinanten Expression eines Fusionsproteins in einem geeigneten Wirt, wobei der Vektor aufweist:

eine Promoterfunktionalität in dem Wirt;

eine DNA gemäß Anspruch 9; und  
eine Stopsequenz.

11. Vektor gemäß Anspruch 10, wobei der Promoter den E. coli- $\beta$ -galactosidase-Promoter umfasst.

12. Vektor gemäß Anspruch 10 oder 11, der ferner einen auswählbaren Marker aufweist.

13. Vektor gemäß einem der Ansprüche 10 bis 12, wobei der Vektor ein Plasmid ist.

14. Vektor gemäß Anspruch 13, der phNF120-1, wie in Beispiel 1, Abschnitt (e) definiert, aufweist.

15. Wirtszelle zur rekombinanten Expression eines Fusionsproteins, wobei die Wirtszelle entweder ein DNA-Konstrukt aufweist, das

eine Promoterfunktionalität in der Wirtszelle;  
eine DNA gemäß Anspruch 9; und  
eine Stopsequenz

aufweist, oder einen Vektor gemäß einem der Ansprüche 10 bis 14.

16. Wirtszelle gemäß Anspruch 15, wobei die DNA oder der Vektor in einem extrachromosalen Element vorliegt.

17. Wirtszelle gemäß Anspruch 15, wobei die DNA oder der Vektor in das Genom der Wirtszelle integriert ist.

18. Verfahren zur Herstellung eines gereinigten Peptids, wobei das Verfahren folgende Schritte aufweist:

Expression eines Fusionsproteins innerhalb einer Wirtszelle gemäß einem der Ansprüche 15 bis 17;

Isolieren des Fusionsproteins von der Wirtszelle endogenen Proteinen durch ein Verfahren, das eine Kationenaustauschchromatographie unter Bedingungen umfasst, bei denen das Fusionsprotein an eine Kationenaustauschsäule gebunden wird, während endogene Proteine der Wirtszelle nicht an die Säule gebunden werden;

Eluieren des Fusionsproteins;

Spalten des Trägerproteins von dem Peptid; und

Trennen des Peptids von dem Trägerprotein.

#### Patentansprüche für folgenden Vertragsstaat : ES

1. Verfahren zur Herstellung eines Fusionsproteins, das vom N-terminalen zum C-terminalen Ende enthält:

ein Trägerprotein von 10 bis 50 kDa, das keine zugänglichen Glu-Reste oder Asp-Gly-Sequenzen als Staph V8-Spaltungsstelle enthält;

eine Staph V8-Spaltungsstelle, die einen Glu-Rest oder eine Asp-Gly-Sequenz aufweist, die an dem C-terminalen Ende des Trägers positioniert ist; und

ein Peptid, das keine als Spaltungsstelle zugänglichen Glu-Reste oder Asp-Gly-Sequenzen enthält, wobei das Peptid an die Spaltungsstelle fusioniert ist und das Fusionsprotein einen pI von 8,0 oder größer zeigt, wobei das Verfahren folgende Schritte umfasst:

Expression des Fusionsproteins innerhalb einer Wirtszelle;

Isolieren des Fusionsproteins von der Wirtszelle endogenen Proteinen durch ein Verfahren, das eine Kationenaustauschchromatographie unter Bedingungen umfasst, bei denen das Fusionsprotein an eine Kat-

ionenaustauschsäule gebunden wird, während endogene Proteine der Wirtszelle nicht an die Säule gebunden werden; und

Eluieren des Fusionsproteins.

2. Verfahren gemäß Anspruch 1, wobei das Trägerprotein im wesentlichen aus der Sequenz besteht:

Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

worin

X<sub>1-9</sub> jeweils unabhängig aus der aus Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, und Tyr bestehenden Gruppe ausgewählt ist; und

X<sub>10</sub> und X<sub>11</sub> jeweils unabhängig aus der aus Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp und Tyr bestehenden Gruppe ausgewählt sind, unter der Voraussetzung, daß X<sub>11</sub> nicht Asp und, wenn X<sub>10</sub> Asp ist, daß X<sub>11</sub> dann nicht Gly ist.

3. Verfahren gemäß Anspruch 2, wobei X<sub>1-9</sub> jeweils Gln ist.
4. Verfahren gemäß einem der Ansprüche 1 bis 3, wobei das Fusionsprotein ferner am N-terminalen Ende eine Leitsequenz von 6 bis 20 Aminosäuren aufweist.
5. Verfahren gemäß Anspruch 4, wobei die Leitsequenz 3 bis 9 Thr-Reste aufweist.
6. Verfahren gemäß Anspruch 5, wobei die Leitsequenz im wesentlichen aus der Sequenz besteht:

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr-Thr-Gln-Phe-  
Arg-Met-.

7. Verfahren gemäß Anspruch 6, wobei das Peptid ein atrionatriuretisches Polypeptid (ANP) ist, ein analoges von ANP, ein Gehirn-natriuretisches Polypeptid, Somatostatin, ein glucagon-ähnliches Peptid, Calcitonin, ein Lungentensid, Insulin, ein Wachstumshormon-Freisetzungsfaktor, ein Bradykinin, ein Endorphin oder ein Enkephalin.
8. Verfahren gemäß einem der Ansprüche 1 bis 7, wobei das Peptid im wesentlichen aus der Sequenz besteht:

Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-Gly-Ala-Gln-  
Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-Tyr.

9. Verfahren zur Herstellung eines Expressionsvektors zur rekombinanten Expression eines Fusionsproteins in einem geeigneten Wirt, wobei der Vektor aufweist:

eine Promotarfunktionalität in dem Wirt;

eine DNA-Sequenz ein Fusionsprotein wie in einen der Ansprüche 1 bis 8 definiert kodierend; und eine Stopsequenz;

wobei das Verfahren die Durchführung der Verbindung des Promoters, der DNA-Kodierungssequenz und der Stopsequenz in einen Vektor umfasst, der die Expression der DNA in einer Wirtszelle steuern kann.

10. Verfahren gemäß Anspruch 9, wobei der Promoter den E. coli- $\beta$ -galactosidase-Promoter umfasst.

11. Verfahren gemäß Anspruch 9 oder 10, wobei der Vektor ferner einen auswählbaren Marker aufweist.

12. Verfahren gemäß einem der Ansprüche 9 bis 11, wobei der Vektor ein Plasmid ist.

13. Verfahren gemäß Anspruch 12, wobei der Vektor phNF120-1, wie in Beispiel 1, Abschnitt (e) definiert, umfasst.

14. Verfahren zur Herstellung einer Wirtszelle zur rekombinanten Expression eines Fusionsproteins, wobei das Verfahren ein Transformieren der Wirtszelle mit entweder einem DNA-Konstrukt umfasst, das

eine Promoterfunktionalität in der Wirtszelle;

ein Fusionsprotein, das eine DNA-Sequenz kodiert wie in einem der Ansprüche 1 bis 8 definiert; und

eine Stopsequenz;

aufweist, oder mit einem Vektor gemäß einem der Ansprüche 9 bis 13.

15. Verfahren gemäß Anspruch 14, wobei die DNA-Sequenz oder der Vektor in einem extrachromosomalen Element vorliegt.

16. Verfahren gemäß Anspruch 14, wobei die DNA-Sequenz oder der Vektor in das Genom der Wirtszelle integriert ist.

17. Verfahren zur Herstellung eines gereinigten Peptids, wobei das Verfahren folgende Schritte aufweist:

Herstellen eines Fusionspeptids, das ein Trägerprotein aufweist, eine Staph V8-Spaltungsstelle und ein Peptid gemäß einem der Ansprüche 1 bis 8;

Spalten des Trägerproteins von dem Peptid; und

Trennen des Peptids vom Trägerprotein.

18. Fusionsprotein, das vom N-terminalen zum C-terminalen Ende aufweist:

ein Trägerprotein von 10 bis 50 kDa, das keine zugänglichen Glu-Reste oder Asp-Gly-Sequenzen als Staph V8-Spaltungsstelle enthält;

eine Staph V8-Spaltungsstelle, enthaltend eine Glu-Rest oder eine Asp-Gly-Sequenz, die an dem C-terminalen Ende des Trägers positioniert ist; und

ein Peptid, das keine als Spaltungsstelle zugängliche Glu-Reste oder Asp-Gly-Sequenzen enthält, das an die Spaltungsstelle fusioniert ist, wobei das Fusionsprotein eine pI von 8,0 oder größer zeigt.

19. DNA, codierend ein Fusionsprotein gemäß Anspruch 18.

20. Expressionsvektor zur rekombinanten Expression, in einem geeigneten Wirt, von einem Fusionsprotein, wobei

der Vektor aufweist:

eine Promoterfunktionalität in dem Wirt;

eine DNA gemäß Anspruch 19; und

eine Terminierungssequenz.

21. Wirtszelle zur rekombinanten Expression eines Fusionsproteins, bei dem die Wirtszelle entweder ein DNA-Konstrukt aufweist mit:

einer Promoterfunktionalität in der Wirtszelle;

eine DNA gemäß Anspruch 19; und

eine Terminierungssequenz;

oder die Wirtszelle einen Vektor gemäß Anspruch 20 aufweist.

# Revendications

Revendications pour les Etats contractants suivants : AT, BE, CH, DE, DK, FR, GB, GR, IT, LI, LU, NL, SE

1. Protéine fusion qui comprend, à partir de l'extrémité N-terminale vers l'extrémité C-terminale :

une protéine porteuse de 10 à 50 kDa qui ne contient pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage de Staph V8 ;

un site de clivage de Staph V8 comprenant un résidu Glu ou une séquence Asp-Gly en position C-terminale dudit porteur ; et

un peptide ne contenant pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage, fusionné audit site de clivage, où ladite protéine de fusion présente un pI de 8,0 ou plus.

2. Protéine de fusion de la revendication 1, dans laquelle ladite protéine porteuse consiste essentiellement en la séquence :

Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

dans laquelle

## EP 0 542 863 B1

X<sub>1-9</sub> sont choisis chacun indépendamment à partir du groupe constitué de Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp et Tyr ; et

X<sub>10</sub> et X<sub>11</sub> sont choisis chacun indépendamment à partir du groupe constitué de Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp et Tyr, à la condition que X<sub>11</sub> ne soit pas Asp et si X<sub>10</sub> est Asp alors X<sub>11</sub> n'est pas Gly.

3. Protéine de fusion de la revendication 2, dans laquelle X<sub>1-9</sub> sont chacun Gln.
4. Protéine de fusion de l'une quelconque des revendications 1 à 3 dans laquelle ladite protéine de fusion comprend en outre une séquence leader N-terminale de 6 à 20 acides aminés.
5. Protéine de fusion de la revendication 4, dans laquelle ladite séquence leader comprend 3 à 9 résidus Thr.
6. Protéine de fusion de la revendication 5, dans laquelle ladite séquence leader consiste essentiellement en la séquence :

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-  
Thr-Gln-Phe-Arg-Met-

7. Protéine de fusion de la revendication 6, dans laquelle ledit peptide est le peptide natriurétique auriculaire (ANP), un analogue de ANP, un peptide natriurétique du cerveau, la somatostatine, un peptide de type glucagon, la calcitonine, le surfactant pulmonaire, l'insuline, le facteur libérant l'hormone de croissance, une bradykinine, une endorphine ou une encéphaline.
8. Protéine de fusion de l'une quelconque des revendications 1 à 7, dans laquelle ledit peptide consiste essentiellement en la séquence :

Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-  
Gly-Ala-Gln-Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-  
Tyr-

9. ADN codant pour une protéine de fusion selon l'une quelconque des revendications 1 à 8.
10. Vecteur d'expression pour l'expression recombinante, dans un hôte approprié, d'une protéine de fusion dans laquelle ledit vecteur comprend :
  - un promoteur fonctionnel dans ledit hôte ;
  - un ADN selon la revendication 9 ; et
  - une séquence de terminaison.
11. Vecteur de la revendication 10, dans lequel ledit promoteur comprend le promoteur de la  $\beta$ -galactosidase de *E. coli*.
12. Vecteur de la revendication 10 ou 11 qui comprend en outre un marqueur qui peut être sélectionné.
13. Vecteur de l'une quelconque des revendications 10 à 12 dans lequel ledit vecteur est un plasmide.
14. Vecteur de la revendication 13, qui comprend pHNF120-1, tel que défini dans l'Exemple 1, section (e).
15. Cellule hôte pour l'expression recombinante d'une protéine de fusion dans laquelle ladite cellule hôte comprend soit une construction d'ADN comprenant :

un promoteur fonctionnel dans ladite cellule hôte ;  
un ADN selon la revendication 9 ; et une séquence de terminaison ;

soit ladite cellule hôte comprend un vecteur selon l'une quelconque des revendications 10 à 14.

16. Cellule hôte de la revendication 15 dans laquelle ledit ADN ou vecteur est présent sur un élément extrachromosomique.

17. Cellule hôte de la revendication 15, dans laquelle ledit ADN ou vecteur est intégré à l'intérieur du génome de ladite cellule hôte.

18. Méthode pour produire un peptide purifié, ladite méthode comprenant les étapes consistant à :

exprimer une protéine de fusion à l'intérieur d'une cellule hôte selon l'une quelconque des revendications 15 à 17 ;

isoler ladite protéine de fusion à partir des protéines endogènes vis-à-vis de ladite cellule hôte par un procédé qui comprend une chromatographie par échange de cations dans des conditions dans lesquelles ladite protéine de fusion est liée à une colonne échangeuse de cations, tandis que les protéines endogènes vis-à-vis desdites cellules hôtes ne sont pas liées à ladite colonne ;

éluer ladite protéine de fusion ;

cliver ladite protéine porteuse à partir dudit peptide ; et  
séparer ledit peptide à partir de ladite protéine porteuse.

#### Revendications pour l'Etat contractant suivant : ES

1. Méthode de production d'une protéine de fusion qui comprend à partir de l'extrémité N-terminale vers l'extrémité C-terminale :

une protéine porteuse de 10 à 50 kDa qui ne contient pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage de Staph V8.

un site de clivage de Staph V8 comprenant un résidu Glu ou une séquence Asp-Gly en position C-terminale dudit porteur ; et

un peptide ne contenant pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage, fusionné audit site de clivage, où ladite protéine de fusion présente un pI de 8,0 ou plus ; méthode qui comprend les étapes consistant à :

exprimer la protéine de fusion à l'intérieur d'une cellule hôte ;

isoler ladite protéine de fusion à partir des protéines endogènes vis-à-vis de ladite cellule hôte par un procédé qui comprend une chromatographie par échange de cations dans des conditions dans lesquelles ladite protéine de fusion est liée à une colonne échangeuse de cations, tandis que les protéines endogènes vis-à-vis desdites cellules hôtes ne sont pas liées à ladite colonne ; et

éluer ladite protéine de fusion.

2. Méthode selon la revendication 1, dans laquelle ladite protéine porteuse consiste essentiellement en la séquence :

Asn-Pro-Met-Tyr-Asn-Ala-Val-Ser-Asn-Ala-Asp-Leu-Met-  
 5 Asp-Phe-Lys-Asn-Leu-Leu-Asp-His-Leu-X<sub>1</sub>-X<sub>2</sub>-Lys-Met-  
 Pro-Leu-X<sub>3</sub>-Asp-X<sub>4</sub>-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-X<sub>5</sub>-  
 Pro-Asn-X<sub>6</sub>-X<sub>7</sub>-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-X<sub>8</sub>-  
 10 Val-Pro-Pro-Trp-Thr-Gly-X<sub>9</sub>-Val-Ser-Pro-Ala-Gln-Arg-X<sub>10</sub>-  
 X<sub>11</sub>-Gly-Ala-Leu-Gly-Arg-Gly-Pro-Trp-Asp-Ser-Ser-Asp-  
 Arg-Ser-Ala-Leu-Leu-Lys-Ser-Lys-Leu-Arg-Ala-Leu-Leu-  
 15 Thr-Ala-Pro-Arg-Ser-Leu-Lys-Phe-;

dans laquelle

20 X<sub>1-9</sub> sont choisis chacun indépendamment à partir du groupe constitué de Ala, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, et Tyr ; et

X<sub>10</sub> et X<sub>11</sub> sont choisis chacun indépendamment à partir du groupe constitué de Ala, Arg, Asp, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, et Tyr, à la condition que X<sub>11</sub> ne soit pas Asp, et si X<sub>10</sub> est Asp alors X<sub>11</sub> n'est pas Gly.

3. Méthode selon la revendication 2, dans laquelle X<sub>1-9</sub> sont chacun Gln.
4. Méthode selon l'une quelconque des revendications 1 à 3 dans laquelle ladite protéine de fusion comprend une  
 30 séquence leader N-terminale de 6 à 20 acides aminés.
5. Méthode selon la revendication 4 dans laquelle ladite séquence leader comprend de 3 à 9 résidus Thr.
6. Méthode selon la revendication 5, dans laquelle ladite séquence leader consiste essentiellement en la séquence :

Met-Thr-Met-Ile-Thr-Asn-Leu-Thr-Thr-Thr-Thr-Thr  
 40 Thr-Gln-Phe-Arg-Met-

7. Méthode selon la revendication 6, dans laquelle ledit peptide est le peptide natriurétique auriculaire (ANP), un  
 45 analogue de ANP, un peptide natriurétique du cerveau, la somatostatine, un peptide de type glucagon, la calcitonine, le surfactant pulmonaire, l'insuline, le facteur libérant l'hormone de croissance, une bradykinine, une endorphine ou une encéphaline
8. Méthode selon l'une quelconque des revendications 1 à 7 dans laquelle ledit peptide consiste essentiellement en la séquence :

Arg-Ser-Ser-Cys-Phe-Gly-Gly-Arg-Met-Asp-Arg-Ile-  
 Gly-Ala-Gln-Ser-Gly-Leu-Gly-Cys-Asn-Ser-Phe-Arg-  
 55 Tyr-

9. Méthode pour la préparation d'un vecteur d'expression pour l'expression recombinante, dans un hôte approprié,



d'une protéine de fusion dans laquelle ledit vecteur comprend :

un promoteur fonctionnel dans ledit hôte ;  
une séquence d'ADN codant pour une protéine de fusion telle que définie dans l'une quelconque des revendications 1 à 8 ; et  
une séquence de terminaison ;  
méthode qui comprend les étapes consistant à ligaturer de manière fonctionnelle ledit promoteur, ladite séquence codante d'ADN et ladite séquence de terminaison dans un vecteur capable de diriger l'expression de la séquence d'ADN dans une cellule hôte.

10. Méthode selon la revendication 9, dans laquelle ledit promoteur comprend le promoteur de la  $\beta$ -galactosidase de *E. coli*.

11. Méthode selon la revendication 9 ou 10 dans laquelle ledit vecteur comprend en outre un marqueur qui peut être sélectionné.

12. Méthode selon l'une quelconque des revendications 9 à 11 dans laquelle ledit vecteur est un plasmide.

13. Méthode selon la revendication 12 dans laquelle ledit vecteur comprend pHNF120-1, tel que défini dans l'Exemple 1, section (e).

14. Méthode pour la préparation d'une cellule hôte pour l'expression recombinante d'une protéine de fusion, méthode qui comprend les étapes consistant à transformer une cellule hôte avec soit une construction d'ADN comprenant :

un promoteur fonctionnel dans ladite cellule hôte ;  
une séquence d'ADN codant pour une protéine de fusion telle que définie dans l'une quelconque des revendications 1 à 8 ; et  
une séquence de terminaison ;  
soit avec un vecteur selon l'une quelconque des revendications 9 à 13.

15. Méthode selon la revendication 14 dans laquelle ladite séquence d'ADN ou ledit vecteur est présent sur un élément extrachromosomique.

16. Méthode selon la revendication 14, dans laquelle ladite séquence d'ADN ou ledit vecteur est intégré l'intérieur du génome de ladite cellule hôte.

17. Méthode pour produire un peptide purifié, ladite méthode comprenant les étapes consistant à produire un peptide de fusion comprenant une protéine porteuse, un site de clivage de Staph V8 et un peptide selon la méthode de l'une quelconque des revendications 1 à 8 ;

cliver ladite protéine porteuse à partir dudit peptide ; et  
séparer ledit peptide de ladite protéine porteuse.

18. Protéine de fusion qui comprend, à partir de l'extrémité N-terminale vers l'extrémité C-terminale:

une protéine porteuse de 10 à 50 kDa qui ne contient pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage de Staph V8;  
un site de clivage de Staph V8 comprenant un résidu Glu ou une séquence Asp-Gly en position C-terminale dudit porteur; et  
un peptide ne contenant pas de résidus Glu ou de séquences Asp-Gly accessibles comme site de clivage, fusionné audit site de clivage, où ladite protéine de fusion présente un pI de 8,0 ou plus.

19. ADN codant pour une protéine de fusion selon la revendication 18.

20. Vecteur d'expression pour l'expression recombinante, dans un hôte approprié, d'une protéine de fusion dans laquelle ledit vecteur comprend :

un promoteur fonctionnel dans ledit hôte;

un ADN selon la revendication 19; et  
une séquence de terminaison.

- 5 21. Cellule hôte pour l'expression recombinante d'une protéine de fusion dans laquelle ladite cellule hôte comprend  
soit une construction d'ADN comprenant :

un promoteur fonctionnel dans ladite cellule hôte;  
un ADN selon la revendication 19; et  
10 une séquence de terminaison;

soit ladite cellule hôte comprend un vecteur selon la revendication 20.

## Staph V8

Met Thr Met Ile Thr Asn Leu Thr Thr Thr Thr Thr Thr Glu Phe Arg  
 CAT ATG ACC ATG ATT ACG AAT TTA ACC ACC ACC ACC ACC GAA TTC CGT  
 CTA TAC TGG TAC TAA TGC TTA AAT TGG TGG TGG TGG TGG CTT AAG GCA

Met Asn Pro Met Tyr Asn Ala Val Ser Asn Ala Asp Leu Met Asp Phe  
 ATG AAC CCG ATG TAC AAC GCA GTG TCC AAC GCA GAC CTG ATG GAT TTC  
 TAC TTG GGC TAC ATG TTG CGT CAC AGG TTG CGT CTG GAC TAC CTA AAG

Lys Asn Leu Leu Asp His Leu Gln Gln Lys Met Pro Leu Gln Asp Gln  
 AAG AAC CTG CTG GAC CAC CTG CAG CAG AAG ATG CCG CTG CAG GAT CAG  
 TTC TTG GAC GAC CTG GTG GAC GTC GTC TTC TAC GGC GAC GTC CTA GTC

Val Val Pro Pro Gln Val Leu Ser Gln Pro Asn Gln Gln Ala Gly Ala  
 GTT GTG CCG CCG CAG GTT CTG TCT CAA CCG AAC CAG CAG GGT GGT GCT  
 CAA CAC GGC GGC GTC CAA GAC AGA GTT GGC TTG GTC GTC CCA CCA CGA

Ala Leu Ser Pro Leu Pro Gln Val Pro Pro Trp Thr Gly Gln Val Ser  
 GCT CTG TCT CCG CTG CCG CAG GTT CCG CCG TGG ACC GGT CAG GTT TCT  
 CGA GAC AGA GGC GAC GGC GTC CAA GGC GGC ACC TGG CCA GTC CAA AGA

Pro Ala Gln Arg Asp Gly Gly Ala Leu Gly Arg Gly Pro Trp Asp Ser  
 CCG GCT CAG CGT GAC GGT GGT GCT CTC GGG CCG GGC CCC TGG GAC TCC  
 GGC CGA GTC GCA CTG CCA CCA CGA GAG CCC GCC CCG GGG ACC CTG AGG

Ser Asp Arg Ser Ala Leu Leu Lys Ser Lys Leu Arg Ala Leu Leu Thr  
 TCT GAT CGA TCT GCC CTC CTA AAA AGC AAG CTG AGG GCG CTG CTC ACT  
 AGA CTA GCT AGA CCG GAG GAT TTT TCG TTC GAC TCC CCG GAC GAG TGA

## Staph V8

Ala Pro Arg Ser Leu Lys Phe Glu Arg Ser Ser Cys Phe Gly Gly Arg  
 CCC CCT CGG AGC CTG AAA TTC GAA CCG TCT TCT TGT TTC GGT GGT CGT  
 CCG GGA GCC TCG GAC TTT AAG CTT GCG AGA AGA ACA AAG CCA CCA GCA

Met Asp Arg Ile Gly Ala Gln Ser Gly Leu Gly Cys Asn Ser Phe Arg  
 ATG GAT CGT ATC GGT GCT CAA TCT GGT TTG GGT TGT AAC TCT TTC AGA  
 TAC CTA GCA TAG CCA CGA GTT AGA CCA AAC CCA ACA TTG AGA AAG TCT

Tyr STP Ala  
 TAC TAA GCT T  
 ATG ATT CGA A

Figure 1

Met Thr Met Ile Thr Asn Leu Thr Thr Thr Thr Thr Thr Gln Phe Arg  
 CAT ATG ACC ATG ATT ACG AAT TTA ACC ACC ACC ACC ACC CAA TTC CGT  
 CTA TAC TGG TAC TAA TGC TTA AAT TGG TGG TGG TGG TGG GTT AAG GCA

Met Asn Pro Met Tyr Asn Ala Val Ser Asn Ala Asp Leu Met Asp Phe  
 ATG AAC CCG ATG TAC AAC GCA GTG TCC AAC GCA GAC CTG ATG GAT TTC  
 TAC TTG GGC TAC ATG TTG CGT CAC AGG TTG CGT CTG GAC TAC CTA AAG

Lys Asn Leu Leu Asp His Leu Gln Gln Lys Met Pro Leu Gln Asp Gln  
 AAG AAC CTG CTG GAC CAC CTG CAG CAG AAG ATG CCG CTG CAG GAT CAG  
 TTC TTG GAC GAC CTG GTG GAC GTC GTC TTC TAC GGC GAC GTC CTA GTC

Val Val Pro Pro Gln Val Leu Ser Gln Pro Asn Gln Gln Ala Gly Ala  
 GTT GTG CCG CCG CAG GTT CTG TCT CAA CCG AAC CAG CAG GCT GGT GCT  
 CAA CAC GGC GGC GTC CAA GAC AGA GTT GGC TTG GTC GTC CGA CCA CGA

Ala Leu Ser Pro Leu Pro Gln Val Pro Pro Trp Thr Gly Gln Val Ser  
 GCT CTG TCT CCG CTG CCG CAG GTT CCG CCG TGG ACC GGT CAG GTT TCT  
 CGA GAC AGA GGC GAC GGC GTC CAA GGC GGC ACC TGG CCA GTC CAA AGA

Pro Ala Gln Arg Asp Gly Gly Ala Leu Gly Arg Gly Pro Trp Asp Ser  
 CCG GCT CAG CGT GAC GGT GGT GCT CTC GGG CCG GGC CCC TGG GAC TCC  
 GGC CGA GTC GCA CTG CCA CCA CGA GAG CCC GCC CCG GGG ACC CTG AGG

Ser Asp Arg Ser Ala Leu Leu Lys Ser Lys Leu Arg Ala Leu Leu Thr  
 TCT GAT CGA TCT GCC CTC CTA AAA AGC AAG CTG AGG GCG CTG CTC ACT  
 AGA CTA GCT AGA CGG GAG GAT TTT TCG TTC GAC TCC CGC GAC GAG TGA

## Staph V8

Ala Pro Arg Ser Leu Lys Phe Glu Arg Ser Ser Cys Phe Gly Gly Arg  
 ACC CCT CGG AGC CTG AAA TTC GAA CGC TCT TCT TGT TTC GGT GGT CGT  
 GGG GGA GCC TCG GAC TTT AAG CTT GCG AGA AGA ACA AAG CCA CCA GCA

Met Asp Arg Ile Gly Ala Gln Ser Gly Leu Gly Cys Asn Ser Phe Arg  
 ATG GAT CGT ATC GGT GCT CAA TCT GGT TTG GGT TGT AAC TCT TTC AGA  
 TAC CTA GCA TAG CCA CGA GTT AGA CCA AAC CCA ACA TTG AGA AAG TCT

Tyr STP Ala  
 TAC TAA GCT T  
 ATG ATT CGA A

Figure 2

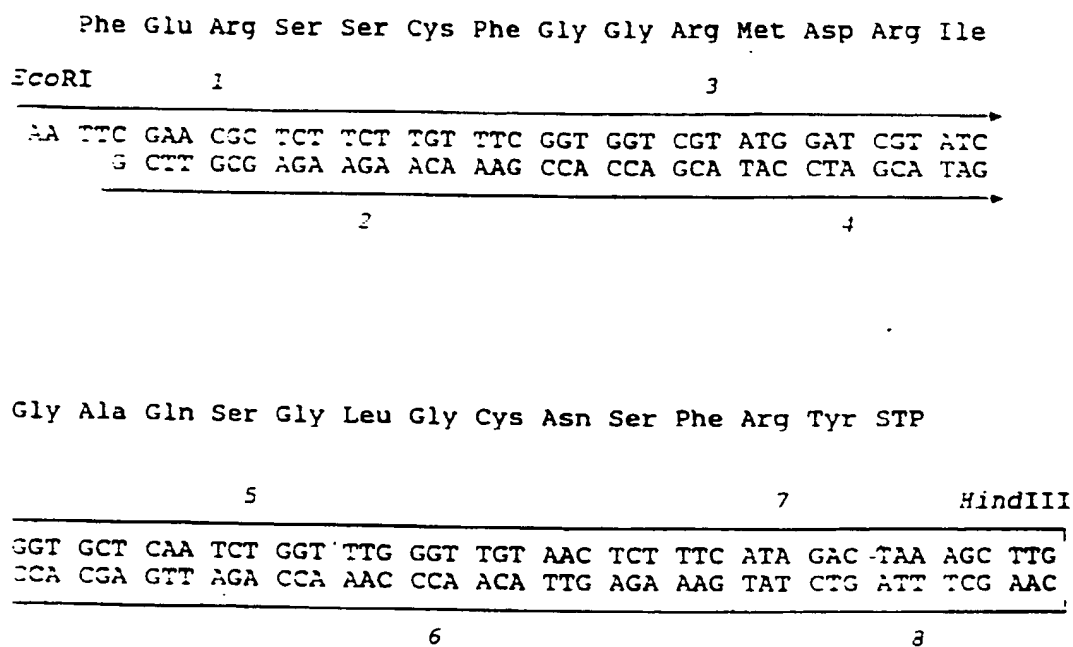


Figure 3

Met Thr Met Ile Thr Asn Leu Thr Thr Thr

NdeI                      :                                  3

T ATG ACC ATG ATT ACG AAT TTA ACC ACC ACC  
AC TGG TAC TAA TGC TTA AAT TGG TGG TGG

2

Thr Thr Glu Phe  
EcoRI                      HindIII

ACC ACC GAA TTC ATT A  
TGG TGG CTT AAG GCA TTC GA

4

### Figure 4

Leu Gly Arg Gly Pro Trp Asp Ser Ser Asp Arg Ser Ala Leu Leu Lys

AvaI

2

3

TC GGG CCG GGC CCC TGG GAC TCC TCT GAT CGA TCT GCC CTC CTA AAA  
C GCC CCG GGG ACC CTG AGG AGA CTA GCT AGA CGG GAG GAT TTT

2

4

Ser Lys Leu Arg Ala Leu Leu Thr Ala Pro Arg Ser Leu Lys Phe

5

7

EcoRI

AGC AAG CTG AGG GCG CTG CTC ACT GCC CCT CGG AGC CTG A  
TCG TTC GAC TCC CGC GAC GAG TGA CGG GGA GCC TCG GAC TTT AA

6

8

Figure 5

Glu Phe Arg Met Asn Pro Met Tyr Asn Ala Val Ser Asn Ala Asp Leu  
 EcoRI 1 3

AA TTC CGT ATG AAC CCG ATG TAC AAC GCA GTG TCC AAC GCA GAC CTG  
 G GCA TAC TTG GGC TAC ATG TTG CGT CAC AGG TTG CGT CTG GAC

2

Met Asp Phe Lys Asn Leu Leu Asp His Leu Gln Gln Lys Met Pro Leu  
5

ATG GAT TTC AAG AAC CTG CTG GAC CAC CTG CAG CAG AAG ATG CCG CTG  
 TAC CTA AAG TTC TTG GAC GAC CTG GTG GAC GTC GTC TTC TAC GGC GAC

4

6

Gln Asp Gln Val Val Pro Pro Gln Val Leu Ser Gln Pro Asn Gln Gln  
7 9

CAG GAT CAG GTT GTG CCG CCG CAG GTT CTG TCT CAA CCG AAC CAG CAG  
 GTC CTA GTC CAA CAC GGC GGC GTC CAA GAC AGA GTT GGC TTG GTC GTC

8

10

Ala Gly Ala Ala Leu Ser Pro Leu Pro Gln Val Pro Pro Trp Thr Gly  
11 13

GCT GGT GCT GCT CTG TCT CCG CTG CCG CAG GTT CCG CCG TGG ACC GGT  
 CGA CCA CGA CGA GAC AGA GGC GAC GGC GTC CAA GGC GGC ACC TGG CCA

12

Gln Val Ser Pro Ala Gln Arg Asp Gly Gly Ala Leu  
15 AvaI

CAG GTT TCT CCG GCT CAG CGT GAC GGT GGT GCT C  
 GTC CAA AGA GGC CGA GTC GCA CTG CCA CCA CGA GAG CC

14

16

# Figure 6



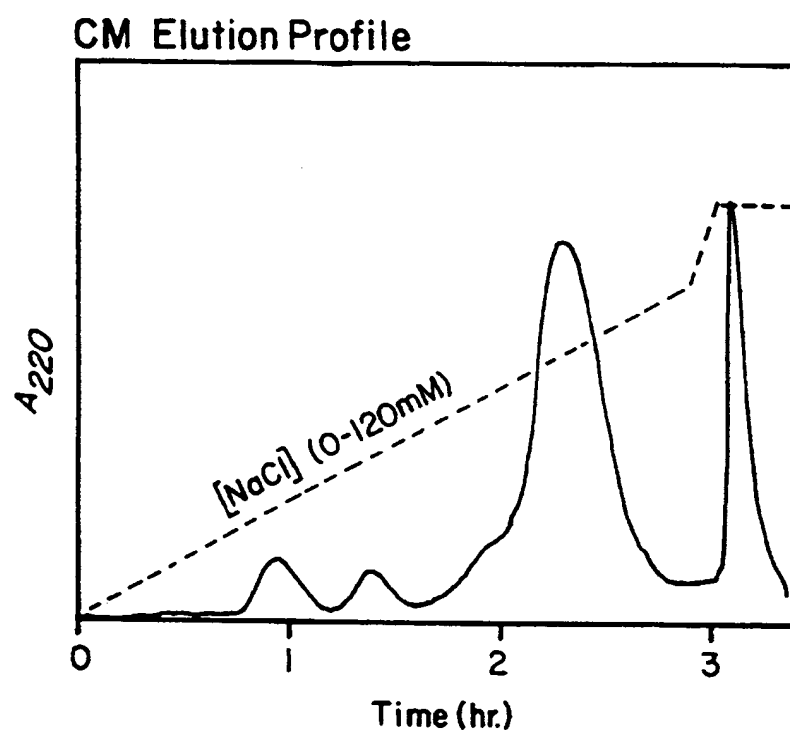


Fig. 7A

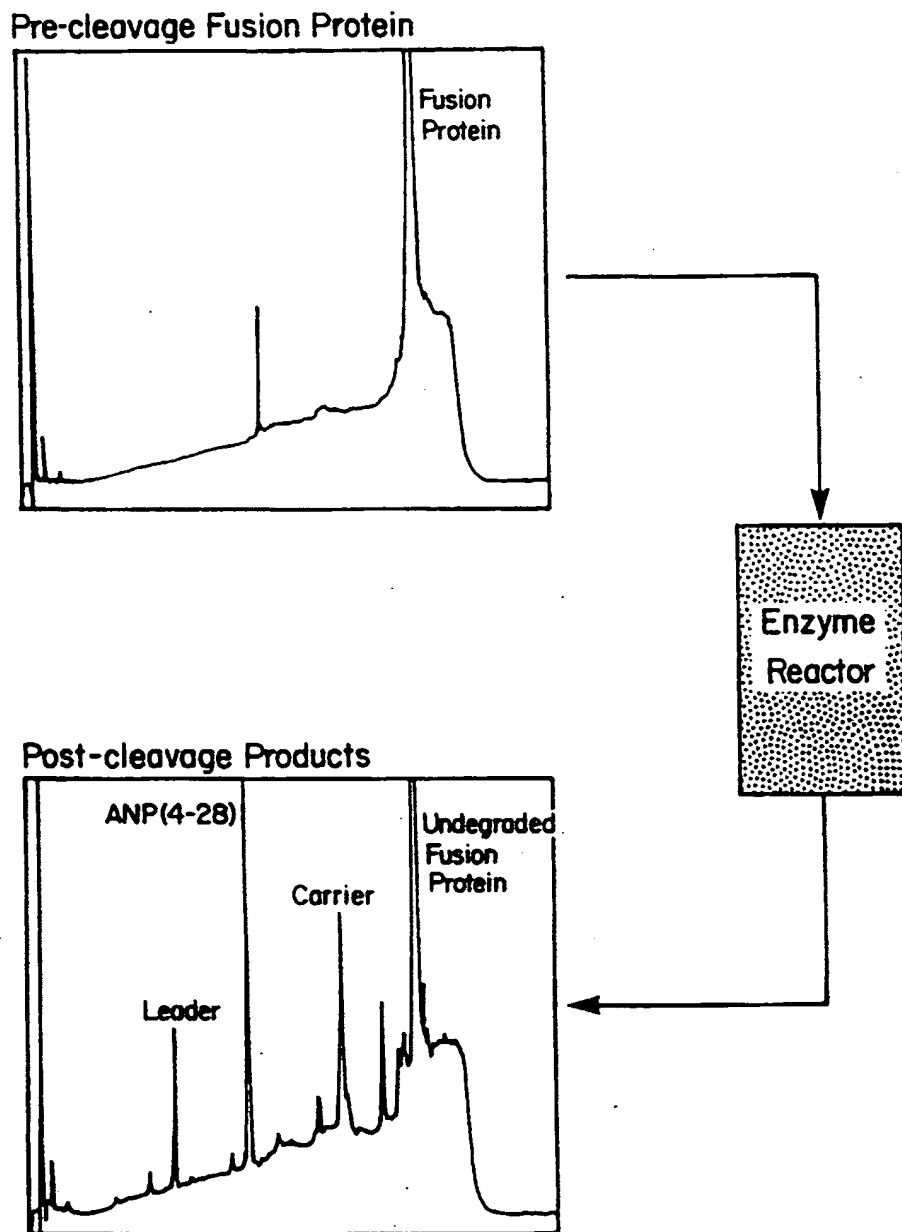
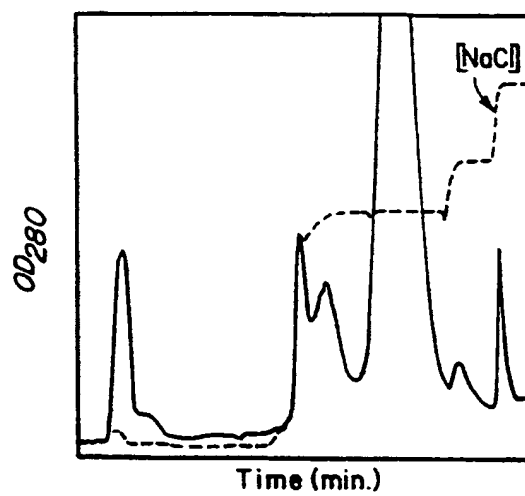


Fig. 7B

CM  
Elution Profile



C-8 RP-HPLC  
Analysis of  
CM Column

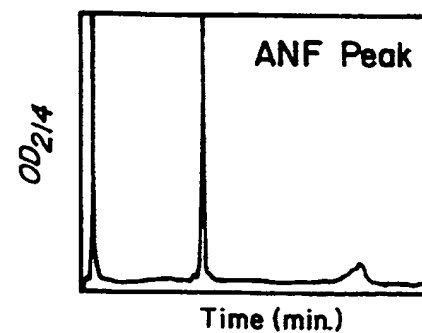
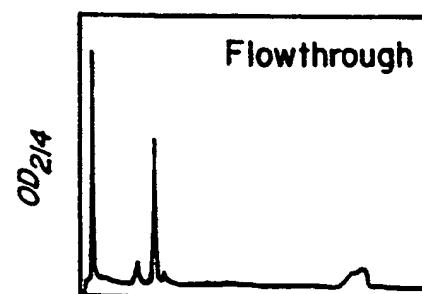
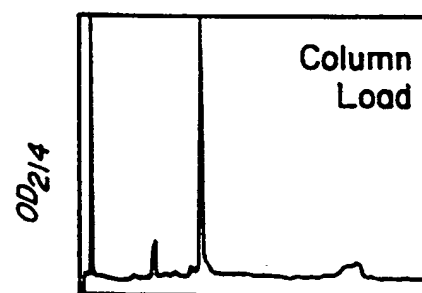


Fig. 7C